



# Motion and Acceleration from Image Assimilation with evolution models

Dominique Béréziat, Isabelle Herlin

## ► To cite this version:

Dominique Béréziat, Isabelle Herlin. Motion and Acceleration from Image Assimilation with evolution models. Digital Signal Processing, 2018, 83, pp.45-58. 10.1016/j.dsp.2018.08.008 . hal-01857811

**HAL Id: hal-01857811**

**<https://hal.sorbonne-universite.fr/hal-01857811>**

Submitted on 14 Nov 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Motion and Acceleration from Image Assimilation with evolution models

Dominique Béréziat (Sorbonne University) \* ; Isabelle Herlin (Inria)

July, 2018

## Abstract

Image sequences allow visualizing dynamic systems and understanding their intrinsic characteristics. One first component of this dynamics is retrieved from the estimation of the velocity displayed on the sequence. Motion estimation has been extensively studied in the literature of image processing and computer vision. In this paper, we step beyond the traditional optical flow methods and address the problem of recovering the acceleration from the whole temporal sequence. This issue has been poorly investigated, even if this is of major importance for major data types, such as fluid flow images. Acceleration is defined as the space-time function resulting from the forces applied to the studied system. To estimate its value, we propose a variational approach where an energy function is designed to model both the motion and the acceleration fields. The contributions of the paper are twofold: first, we introduce a unified variational formulation of motion and acceleration under space-time constraints; second, we describe the minimization scheme, which allows retrieving the estimations, and provide the full information on the discretization schemes. Last, experiments illustrate the potentiality of the method on synthetic and real image sequences, visualizing fluid-like flows, where direct and precise calculation of acceleration is of primary importance.

## Keywords

acceleration, Coriolis force, data assimilation fluid flow, gravity force, motion, shallow water

## 1 Introduction

Estimating motion from images is of major importance for a large range of environmental applications. Analyzing satellite acquisitions of Sea Surface Temperature allows, for instance, to detect precursors of extreme events and better

---

\*<mailto:dominique.bereziat@lip6.fr>, corresponding author

mitigate their risks. Processing fish-eye sky images acquired on a solar plant allows forecasting the solar irradiance and accurately estimating the photovoltaic production.

Motion estimation on fluid flows has been extensively discussed in, for instance, Heitz *et al.* [1]. The underlying problems of the fluid flows context are quite different of those that are usually occurring in most computer vision applications. One major difference is that these fluid flows data require a dense in space and dense in time description and can not be summarized by local features, which are tracked in time. This is the spirit of the research work described in the following and it explains most of the technical choices that were done when implementing the approach.

Acceleration estimation seems having been seldom investigated in the state-of-the-art, whatever the type of image acquisitions. In Arnspang [2], the assumption of Lagrangian constancy of both brightness and velocity is used to estimate motion and acceleration on three consecutive frames. However, these fields are supposed to be locally constant. Hu and Ahuja [3] propose an approach, which is not based on the image data but on a set of characteristics points tracked on consecutive acquisitions, and compute the affine and projective parameters of motion and a constant, in space and time, acceleration. Staković *et al.* [4] determine the motion and acceleration fields using Fourier-based techniques, but motion is restricted to be locally translational and acceleration is also locally constant.

Multi-frame motion has been the focus of a large number of research works. Tomasi and Kanade [5] and Irani [6] extend the pioneer work of Lucas and Kanade [7] in a multi-frame context. Under the assumption of locally stationary motion, they prove that motion fields are included in a low rank subspace. The main advantage of their methods is the robustness to noise and the capability to solve the aperture problem without applying any regularization process. In the same spirit, Garg *et al.* [8] also compute non rigid motion fields from their projection on a low rank subspace. Ricco and Tomasi [9] define a multi-frame method to assess long-range motion and detect the occlusions by computing the Lagrangian trajectories of points on the low rank subspace. Garg *et al.* [10] apply the same concept for image registration purpose. The major limitation of this whole set of these methods, compared with our approach, is the poor representation of the dynamics, with a small size basis, which is precise enough for their application domains but fails for the complex motion fields of fluid flows. This remark originates our motivation for implementing a dynamic model  $\mathbf{M}$  describing both the temporal evolution of the motion field and the transport of the image data.

Accuracy of motion estimation, both in direction and intensity is also a key component for forecasting images at short temporal horizon and mitigating future events. This concern is particularly crucial for environmental issues as, for instance, the short-term forecast of heavy rain or clear sky, which are required for mitigating flashfloods or estimating and regulating the production of photovoltaic energy. However, the pertinence of the forecasted images relies on the full knowledge of the space-time dynamics, and not only on the motion field

component. In the operational application which is considered for the discussion given in the paper, such forecast is applied with a temporal sliding-window setting: a set of images is first processed for estimating motion and acceleration, which are further used for forecasting the image data at a given temporal horizon, then the temporal window is iteratively incremented in time. In such context, the only knowledge of motion snapshots is not sufficient for a correct forecast of the future motion fields and a dense in time model of motion and acceleration is strongly required.

For a large range of environmental systems observed with image data, mathematical models of the physical processes are available. This is well-known for meteorology and oceanography for instance, which are both based on the Navier-Stokes equations. These physical laws should then be used when processing the image acquisitions, in order to allow a full and reliable estimation of the dynamics. The paper describes the design of an image model that includes evolution equations for all studied quantities such as image brightness, velocity and acceleration. Then, we discuss the estimation of the full dynamics, motion and acceleration, with a data assimilation approach, which originates in the meteorology forecasting community and is currently used in meteorological institutes all over the world. These data assimilation techniques appeared in the last decade in the image processing and computer vision community for estimating motion from image sequences, as for instance in Papadakis *et al.* [11, 12], Titaud *et al.* [13] and Béréziat *et al.* [14]. One primary output of these approaches is an elegant solution of the well-known aperture ambiguity by an explicit model of motion. But the paper makes a strong improvement, compared to these state-of-the-art methods, as it allows the estimation of forces applied to the system, or equivalently the simultaneous estimation of motion and acceleration. For an accurate comparison with the literature, we highlight that Papadakis *et al.* [11, 12] or Heas *et al.* [15] also include an additional quantity, which could be viewed as an acceleration term. However, their mathematical formulation constrains this quantity to be small or sparse, consequently suppressing any physical interpretation. In this paper, we focus on a class of approaches named 4D-Var in the data assimilation literature that are relying on an adjoint formulation. However ensemble-based approaches are also possible [16]. Our approach solves the following inverse problem: given  $N^O$  images  $I_l^O$ ,  $l = 1, \dots, N^O$ , the motion and acceleration fields are estimated under the constraint of the given dynamic model, expressed by partial differential equations, and space-time regularity properties. Compared to the previously mentioned motion estimation methods that rely on data assimilation, an additional equation is added to the model, which corresponds to the description of the acceleration. If this equation includes a parametric formulation of the acceleration [17], the problem reduces to the estimation of the parameters values without any strong difficulty. But, in the general case of fluid-flows images, the parametric assumption is not valid on the data and a variational data assimilation technique is applied for estimating a dense acceleration field [18]. A specific energy is then designed whose control variables are the values of all variables at the beginning of the studied temporal interval and the acceleration field at each space-time value. The optimization is

conducted by computing iteratively the values of the energy and of its gradient, which are the input of the optimization solver, named BFGS [19]. The outputs are the motion and acceleration fields on a continuous temporal interval.

*Paper Organization:* Section 2 discusses the problem faced in the paper, provides the basic notations and describes the mathematical content. The variational data assimilation technique is then shortly discussed in Section 3, which provides the main mathematical and technical components for understanding the approach. Our method is extensively described in Section 4 for allowing interested Readers to reproduce the full implementation and the experiments. Results are thoroughly discussed in Section 5, and Section 6 concludes the paper and gives indications on future work.

## 2 Mathematical setting

In order to improve the understanding of our approach, all symbols included in the paper are the same for continuous and discrete descriptions, even if not always fully correct from the mathematical point of view.

$\Omega$  is the image domain and  $[0, T]$  is the temporal interval on which images are acquired and processed. The set  $A = \Omega \times [0, T]$  is the studied space-time domain. If a function  $f$  is defined on  $A$ ,  $f(\mathbf{x}, t)$  denotes the value at point  $\mathbf{x}$  and time  $t$  and  $f(t)$  describes the spatial field at time  $t$ .

The motion vector at point  $\mathbf{x}$  and time  $t$  is written  $\mathbf{w}(\mathbf{x}, t) = (u(\mathbf{x}, t) \ v(\mathbf{x}, t))^T$  with  $^T$  being the transpose operator and  $u$  and  $v$  the horizontal and vertical components. The acceleration is written  $\mathbf{a}(\mathbf{x}, t) = (a_u(\mathbf{x}, t) \ a_v(\mathbf{x}, t))^T$ .

$\langle f, g \rangle$  denotes the scalar product of functions  $f$  and  $g$  in the continuous domain (or in the discrete domain) and verifies:

$$\langle f, g \rangle = \int_{\Omega} f(x)g(x)dx \quad (1)$$

A discrete sequence of images  $I_l^O$ ,  $l = 1, \dots, N^O$ , is available and processed for estimating motion and acceleration.  $I_l^O$  is acquired at time  $t_l$  and is a snapshot of the continuous function  $I^O$ , defined on  $A$ , with values  $I^O(\mathbf{x}, t_l)$ .

As pointed out in the introduction, our approach estimates motion and acceleration from images and is based on a dynamic model  $\mathbf{M}$ . A state vector  $\mathbf{X}$  is first defined on  $A$ :  $\mathbf{X}(\mathbf{x}, t) = (u(\mathbf{x}, t) \ v(\mathbf{x}, t) \ I(\mathbf{x}, t))^T$ , which includes the two components  $u$  and  $v$  of the velocity  $\mathbf{w}$  and a synthetic image  $I$ . The function  $I$  satisfies the same physical and mathematical properties than the real image acquisitions. It is initialized with the first image  $I_1^O$ , at the beginning of the studied temporal interval, and transported by the motion field  $\mathbf{w}(\mathbf{x}, t)$ . If this transport is correctly performed by the estimated motion field, the brightness values  $I(\mathbf{x}, t)$  should be almost identical to the image values  $I_l^O(\mathbf{x})$  at each acquisition time  $t_l$ . Consequently, our method estimates motion and acceleration by forcing the image function  $I$  to be almost identical to the observed images  $I_l^O$ .

The model  $\mathbb{M}$ , expressing the evolution of the state vector, is defined by partial differential equations regulating the time evolution of  $\mathbf{w}$ ,  $\mathbf{a}$  and  $I$ . The motion and acceleration functions are mathematically linked by:

$$\frac{d\mathbf{w}}{dt}(\mathbf{x}, t) = \frac{\partial \mathbf{w}}{\partial t}(\mathbf{x}, t) + (\mathbf{w} \cdot \nabla) \mathbf{w}(\mathbf{x}, t) = \mathbf{a}(\mathbf{x}, t) \quad (2)$$

The evolution equation chosen for the synthetic image function  $I$  is the linearized transport equation:

$$\frac{\partial I(\mathbf{x}, t)}{\partial t} + \mathbf{w}(\mathbf{x}, t) \cdot \nabla I(\mathbf{x}, t) = 0 \quad (3)$$

In the following, the symbol  $\mathbf{a}$  denotes both the acceleration function, defined on  $A$ , and its representation in the state vector space:  $\mathbf{a} = \begin{pmatrix} \mathbf{a} \\ 0 \end{pmatrix}$ . Eqs. (2) and (3) are then merged in:

$$\frac{\partial \mathbf{X}}{\partial t} + \mathbb{M}(\mathbf{X}) = \mathbf{a} \quad (4)$$

The optimization process, that will be described in the next section, constrains the synthetic image function  $I$  to be close to the observed image function  $I^O$ . A measure of the reconstruction error is given by:

$$I(\mathbf{x}, t) = I^O(\mathbf{x}, t) + \varepsilon_R(\mathbf{x}, t) \quad (5)$$

where the term  $\varepsilon_R(\mathbf{x}, t)$  characterizes both the acquisition noise of real images and the uncertainty on the state vector values  $\mathbf{X}(\mathbf{x}, t)$ .

At each iteration of the optimization estimating motion and acceleration, Eq. (4) is integrated in time from an initial value  $\mathbf{X}_b(\mathbf{x})$ . The result of the optimization,  $\mathbf{X}(\mathbf{x}, 0)$ , differs from the value  $\mathbf{X}_b(\mathbf{x})$  as this process modifies the state vector according to the image acquisitions. This idea is described by the following equation:

$$\mathbf{X}(\mathbf{x}, 0) = \mathbf{X}_b(\mathbf{x}) + \varepsilon_B(\mathbf{x}) \quad (6)$$

where  $\varepsilon_B(\mathbf{x})$  denotes the uncertainty on the initial value  $\mathbf{X}_b(\mathbf{x})$ .

The error functions  $\varepsilon_R$  and  $\varepsilon_B$  are supposed independent, unbiased and Gaussian.

For estimating the motion field  $\mathbf{w}$  and the acceleration  $\mathbf{a}$  from images, according to Eqs. (4), (5) and (6), the variables  $\varepsilon_B(\mathbf{x})$  (or  $\mathbf{X}(\mathbf{x}, 0) - \mathbf{X}_b(\mathbf{x})$ ) and  $\varepsilon_R(\mathbf{x}, t)$  (or  $I(\mathbf{x}, t) - I^O(\mathbf{x}, t)$ ) are minimized as described in the next section, dedicated to the data assimilation theory.

### 3 Data Assimilation

The previous equations, Eqs (4, 5, 6), model the links between the observational images and the mathematical model  $\mathbb{M}$ . They are the base of the so-called data assimilation approach. Apte *et al.* [20] gives a Bayesian view of data assimilation, in which the posterior distribution of  $\mathbf{X}$  is computed from the prior distributions of the errors  $\varepsilon_B$  and  $\varepsilon_R$ . Their paper defines various optimization

approaches, which output *a posteriori* estimators. Following their description, we are looking for the solution  $\mathbf{X}$  that solves System (4,5,6) and minimizes the error terms on the temporal interval  $[0, T]$ . An energy function  $J$  is defined, whose control variables are the initial value  $\mathbf{X}(0)$  and the acceleration field  $\mathbf{a}(t)$  at each time  $t$ . This function is primary written as the squared norm of the two error variables  $\varepsilon_B$  and  $\varepsilon_R$  or equivalently of  $(\mathbf{X}(0) - \mathbf{X}_b)$  and  $(I(t) - I^O(t))$ :

$$2J(\mathbf{X}(0), \mathbf{a}(t)) = \langle \mathbf{X}(0) - \mathbf{X}_b, B^{-1}(\mathbf{X}(0) - \mathbf{X}_b) \rangle + \int_0^T \langle I(t) - I^O(t), R^{-1}(I(t) - I^O(t)) \rangle \quad (7)$$

The first term issues from Eq. (6) on the initial condition and the second from Eq. (5) on the image acquisitions.  $B$  and  $R$  denote the covariance matrices of  $\varepsilon_B$  and  $\varepsilon_R$ .

The gradient of  $J$ , denoted  $\nabla J$ , is derived with the well-known calculus of variations, as initially defined by Lions [21]. Its two components, which are respectively the partial derivatives with regard to  $\mathbf{X}(0)$  and  $\mathbf{a}(t)$ , are given by:

$$\begin{aligned} \frac{\partial J}{\partial \mathbf{X}(0)} &= B^{-1}(\mathbf{X}(0) - \mathbf{X}_b) + \lambda(0) \\ \frac{\partial J}{\partial \mathbf{a}(t)} &= \lambda(t) \end{aligned} \quad (8)$$

$\lambda(t)$  is the adjoint variable of  $\mathbf{X}(t)$ , which is mathematically defined by:

$$\lambda(T) = 0 \quad (9a)$$

$$-\frac{\partial \lambda(t)}{\partial t} + \left( \frac{\partial \mathbf{M}}{\partial \mathbf{X}} \right)^* \lambda(t) = R^{-1}(I(t) - I^O(t)) \quad (9b)$$

where  $\left( \frac{\partial \mathbf{M}}{\partial \mathbf{X}} \right)^*$  is the adjoint model of  $\mathbf{M}$ .  $\lambda(t)$  is computed from its initial value  $\lambda(T)$  with a backward in time integration of the the adjoint model. Reader is referred to [22] for a full proof.

The optimization minimizes the error between the given image sequence and the synthesized image sequence through the minimization of the energy  $J(\mathbf{X}(0), \mathbf{a}(t))$  with an iterative steepest descent method. In our implementation, we make use of the low memory BFGS solver [19]. At each iteration, the forward integration of  $\mathbf{X}(t)$ , according to Eq. (4), provides the value of  $J$ , and the backward integration of  $\lambda(t)$ , according to Eq. (9b), computes the values of  $\lambda(0)$  and  $\lambda(t)$ , which are used to compute  $\nabla J$ , as specified by Eq. (8).

The next section describes the use of these data assimilation tools in the context of motion and acceleration estimation from observational images and gives full knowledge of the additional components.

## 4 Motion and acceleration estimation

The implementation of the motion and acceleration estimator is obtained with a numerical model discretizing the continuous equations, Eqs. (2) and (3), of the image model  $\mathbf{M}$ . Time and space resolutions are denoted by  $\Delta t$  and  $\Delta x$  (which is equal to  $\Delta y$ ), and  $u_{i,j}^n$  gives the approximation of  $u(x, y, t)$  at location  $(x_i, y_j, t_n)$ . The equations are first discretized in time with an Euler scheme. The linear advection terms are discretized in space with a first-order upwind scheme, as described in Hundsdorfer and Spee [23]. The nonlinear terms are rewritten in a conservative form and approximated by a first-order Godunov scheme, given in Leveque [24]. For instance, the non-linear advection of the horizontal component of velocity is rewritten as:

$$\frac{\partial u}{\partial t} = -u \frac{\partial u}{\partial x} = -\frac{\partial}{\partial x} f(u) \quad (10)$$

with  $f(u) = \frac{1}{2}u^2$ . This equation is then approximated by:

$$u_{i,j}^{n+1} = u_{i,j}^n - \frac{\Delta t}{\Delta x} (g(u_{i,j}^n, u_{i+1,j}^n) - g(u_{i-1,j}^n, u_{i,j}^n)) \quad (11)$$

with  $g$  defined by:

$$\begin{aligned} \text{if } s_1 < s_2 : \quad g(s_1, s_2) &= \begin{cases} f(s_2) & \text{if } s_2 \leq 0 \\ f(s_1) & \text{if } s_1 \geq 0 \\ 0 & \text{otherwise} \end{cases} \\ \text{if } s_1 \geq s_2 : \quad g(s_1, s_2) &= \begin{cases} f(s_2) & \text{if } s_1 \leq 0 \\ f(s_1) & \text{if } s_2 \geq 0 \end{cases} \end{aligned} \quad (12)$$

The backward in time integration of the adjoint variable  $\lambda$  involves the adjoint operator  $\left(\frac{\partial \mathbf{M}}{\partial X}\right)^*$ , in Eq. (9b), that is obtained by the automatic differentiation software Tapenade [25] from the numerical model  $\mathbf{M}$ . The numerical adjoint obtained with Tapenade is the exact adjoint of  $\mathbf{M}$ , as each line of the numerical code is differentiated by the software.

Having coded the direct model and derived its numerical adjoint, the implementation of our approach is based on the definition of the following components: the initial condition  $\mathbf{X}_b$ , the characterization of  $\varepsilon_B$  and  $\varepsilon_R$ , the design of the regularization terms, and the implementation of the minimization process.

- The covariance matrix  $B$  models the uncertainty on the initial value  $\mathbf{X}_b$ . We assume that the three components of  $\mathbf{X}_b$  are not correlated and set

$$B = \begin{pmatrix} B_u & 0 & 0 \\ 0 & B_v & 0 \\ 0 & 0 & B_I \end{pmatrix}.$$

- The initial value  $\mathbf{X}_b = (u_b \quad v_b \quad I_b)^T$  is set to  $(0 \quad 0 \quad I_1^O)^T$ . The image field  $I_b$  is given the value of the first acquisition  $I_1^O$ , as the image function



$I$  has to be close to the observations for a correct estimation of motion and acceleration. The motion field  $\mathbf{w}_b = (u_b \ v_b)^T$  is given a null value at the beginning of the optimization process if no prior information is available. We then give an infinite uncertainty to  $u_b$  and  $v_b$  and set  $B_u^{-1} = B_v^{-1} = 0$ . Consequently, the first term of  $J$  reduces to  $\langle I(0) - I_b, B_I^{-1}(I(0) - I_b) \rangle$ . The reader should remark that the initial motion field has no constraint to stay close to this null value in the estimation process with such value of the covariance matrix  $B$ .

The alternative, which is the one used in our operational setting, relies on a sliding-window approach: - the initial value of the motion field of the first window is again given a null value, - but, for the next windows, the initial value is defined as the estimation on the previous one. Again,  $B_u^{-1}$  and  $B_v^{-1}$  are given a zero value as  $u_b$  and  $v_b$  being not observed, their uncertainty can not be quantified.

- The covariance matrix  $B_I$  of the error term  $\varepsilon_B(\mathbf{x})$  depends on the location  $\mathbf{x}$ . It describes the uncertainty on the initial image value  $I_1^O(\mathbf{x})$ . It is chosen equal to that of the acquisition sensor. However, noisy pixels (identified as such from the metadata provided with images) get the value  $B_I(\mathbf{x}) = 10^8$  so that their contribution in the energy function  $J$  of Eq. (7) is infinitesimal.
- The covariance matrix  $R$  of the error term  $\varepsilon_R(\mathbf{x}, t)$  depends on  $\mathbf{x}$  and  $t$  and describes the uncertainty on image acquisitions.  $R(\mathbf{x}, t)$  gets the same value than  $B_I(\mathbf{x})$ . The same process is applied to discard the noisy pixels from the computation.
- The estimation of  $\mathbf{w}(0)$  and  $\mathbf{a}(t)$  from a sequence of images remains an ill-posed problem. As we assume that the initial motion field is Markovian Gaussian rather than degenerated Gaussian, the cost function  $J$  includes spatial regularization terms, which are depending on the norm of the gradient of  $\mathbf{w}(0)$  and  $\mathbf{a}(t)$ .

$$\begin{aligned}
2J(\mathbf{X}(0), \mathbf{a}(t)) &= \langle I(0) - I_b, B_I^{-1}(I(0) - I_b) \rangle \\
&+ \int_0^T \langle I(t) - I^O(t), R^{-1}(I(t) - I^O(t)) \rangle \\
&+ \alpha \langle \nabla \mathbf{w}(0), \nabla \mathbf{w}(0) \rangle + \beta \int_0^T \langle \nabla \mathbf{a}(t), \nabla \mathbf{a}(t) \rangle dt
\end{aligned} \tag{13}$$

The gradient of  $J$  is:

$$\begin{aligned}
\frac{\partial J}{\partial \mathbf{X}(0)} &= B_I^{-1}(I(0) - I_b) - \alpha \nabla^2 \mathbf{w}(0) + \lambda(0) \\
\frac{\partial J}{\partial \mathbf{a}(t)} &= -\beta \nabla^2 \mathbf{a}(t) + \lambda(t)
\end{aligned} \tag{14}$$

$\lambda$  remains ruled by Equations (9a,9b).

Values of  $\alpha$  and  $\beta$  are chosen so that the regularization terms get the same range order than the background term.

As previously explained, the minimization of the energy  $J$  is conducted with a steepest descent method and the BFGS solver. The full description is given by Algorithm 1, in case of null initial values of motion and acceleration values.

---

**Algorithm 1** 4D-Var algorithm.

---

- 1: Read the observational images  $I_l^O$ ,  $l = 1, \dots, N^O$
  - 2: Set the initial background value  $\mathbf{X}_b = (0 \ 0 \ I_1^O)^T$
  - 3: Set the iteration index  $k = 0$
  - 4: Initialize the state vector value  $\mathbf{X}(0)^k = \mathbf{X}_b$
  - 5: Set the initial acceleration value:  $\mathbf{a}^k = (0 \ 0)^T$
  - 6: Run the model IM (forward integration) and compute  $J^k$
  - 7: Run (backward) the adjoint model and compute  $\nabla J^k$
  - 8: **while**  $|\nabla J^k| > \mathcal{E}$  **and**  $k < \text{MaxIter}$  **do**
  - 9:     Update state vector  $\mathbf{X}(0)$  and acceleration  $\mathbf{a}$ :  
 $(\mathbf{X}(0)^{k+1}, \mathbf{a}^{k+1}) = \text{BFGS}(\mathbf{X}(0)^k, \mathbf{a}^k, J^k, \nabla J^k)$
  - 10:     Run the model IM and compute  $J^{k+1}$
  - 11:     Run the adjoint model and compute  $\nabla J^{k+1}$
  - 12:     Set  $k = k + 1$
  - 13: **end while**
  - 14: Set  $k = k - 1$
  - 15: Return  $\mathbf{X}(0)^k$  and  $\mathbf{a}^k(t)$  for each time index  $t$
- 

At this stage, all components of our approach have been discussed. Additional setting of the implementation are depending on the application domain and on the operational use of images. They will be discussed jointly with the analysis of the results in the next section.

## 5 Results

As explained in the previous discussions, our approach is a strong theoretical improvement compared to the state-of-the-art on motion estimation with data assimilation techniques and to our own research work on the subject [26]. Having knowledge of the acceleration gives a full understanding of the observed system and allows, for instance, to accurately forecast the future images.

This section demonstrates the impact of our approach on fluid flow data and provides results on synthetic and satellite image sequences.

### 5.1 Synthetic Experiments

The method is first applied on a synthetic sequence in order to get a quantitative evaluation. The image and motion fields are obtained by integrating Eqs. (2,3)

with the initial conditions displayed on Figure 1. The acceleration  $\mathbf{a}(\mathbf{x}, t)$  is

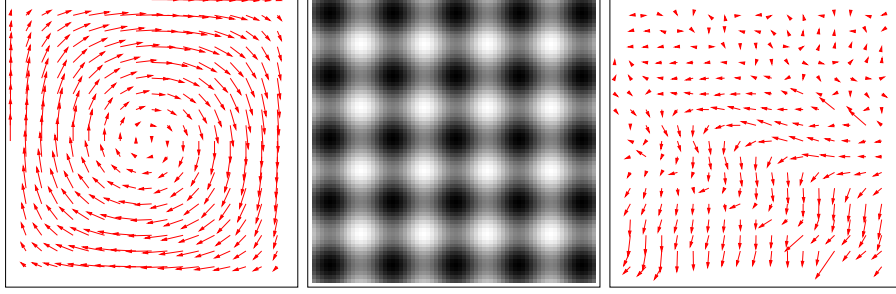


Figure 1: Initial conditions (left and middle) and acceleration term (right).

chosen to be constant in time, consequently denoted  $\mathbf{a}(\mathbf{x})$ , and modeled by a spatial Balgovind function [27],  $\mathbf{a}(\mathbf{x}) = \begin{pmatrix} \mathcal{B}_u(\mathbf{x}) \\ \mathcal{B}_v(\mathbf{x}) \end{pmatrix}$ , which is visualized on the right of Figure 1.  $\mathcal{B}_u(\mathbf{x})$  and  $\mathcal{B}_v(\mathbf{x})$  are zero-mean normal laws, whose covariance values depend on the distance between  $\mathbf{x}$  and  $\mathbf{x}'$ :

$$\Sigma(\mathbf{x}, \mathbf{x}') = \left(1 + \frac{\|\mathbf{x} - \mathbf{x}'\|}{L}\right) \exp\left(-\frac{\|\mathbf{x} - \mathbf{x}'\|}{L}\right) \quad (15)$$

$L$  is the bandwidth:  $\mathbf{x}$  and  $\mathbf{x}'$  are statistically independent if  $\|\mathbf{x} - \mathbf{x}'\| > L$ . In this experiment,  $L$  is set to 20% of the image size.

Five snapshots of the image component  $I(t)$  are taken as observational images. They are displayed on Figure 2. The method estimates motion and acceleration on the whole time interval from these observations. In order to assess the performance, we compare the results on motion and acceleration with the ground truth values, obtained from the simulation run, and we compute errors statistics. The averaged angular and relative norm errors on the motion field are respectively of 4.5 degrees and 7%. This demonstrates the accuracy of both motion and acceleration estimation, as motion can be correctly retrieved on the whole time interval only if acceleration is simultaneously accurately assessed.

In a second experiment, we show that our approach is able to retrieve motion and acceleration on synthetic data having the same characteristics than satellite acquisitions of the ocean surface. We assume that the surface circulation is described by the shallow water equations, which characterize the temporal evolution of the motion field by the gravity and Coriolis forces, and neglect the surface friction and the diffusion in this experiment.

$$\begin{cases} \frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} + v \frac{\partial u}{\partial y} &= f v - g' \frac{\partial h}{\partial x} \\ \frac{\partial v}{\partial t} + u \frac{\partial v}{\partial x} + v \frac{\partial v}{\partial y} &= -f u - g' \frac{\partial h}{\partial y} \\ \frac{\partial h}{\partial t} + \frac{\partial(hu)}{\partial x} + \frac{\partial(vh)}{\partial y} &= 0 \end{cases} \quad (16)$$

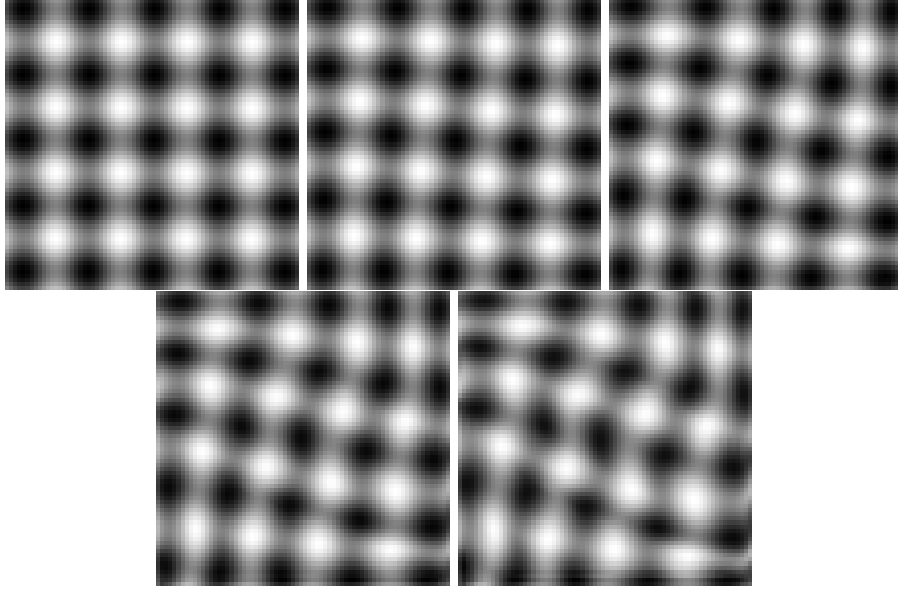


Figure 2: Observational images.

In System (16),  $h$  is the mixed layer thickness,  $f$  the Coriolis parameter, and  $g' = g(\rho_0 - \rho_1)/\rho_0$  the reduced gravity,  $\rho_0$  being the reference density and  $\rho_1$  the average density of the mixed layer.

We further summarize the geophysical forces, gravity and Coriolis, as an acceleration term  $\mathbf{a}$  with two components  $\mathbf{a}_u$  and  $\mathbf{a}_v$ :

$$\begin{cases} \mathbf{a}_u &= f v - g' \frac{\partial h}{\partial x} \\ \mathbf{a}_v &= -f u - g' \frac{\partial h}{\partial y} \end{cases} \quad (17)$$

The synthetic experiment is then designed. Given the initial conditions on velocity  $\mathbf{w}$ , layer thickness  $h$ , and synthetic image  $I$  visualized on Figure (3), the shallow-water equations of System (16) and the transport of the image function with Eq. (3) are integrated in time. These initial conditions come from experiments conducted on ocean satellite images that are not further discussed in the paper.

The numerical simulation provides eight image snapshots that are used as observational data for our motion and acceleration approach. Our method retrieves motion with an averaged angular error of  $5.3^\circ$  and an averaged relative norm error of 11%. As a comparison, Sun *et al.* [28] obtain, on the same data, an averaged angular error of  $16^\circ$  and a relative norm error of 32%. This comes from the fact that their model does not include information on the physical forces or equivalently on the acceleration. Additionally, the estimation of velocity  $\mathbf{w}$  and acceleration  $\mathbf{a}$  allows us to further derive the Coriolis and gravity forces

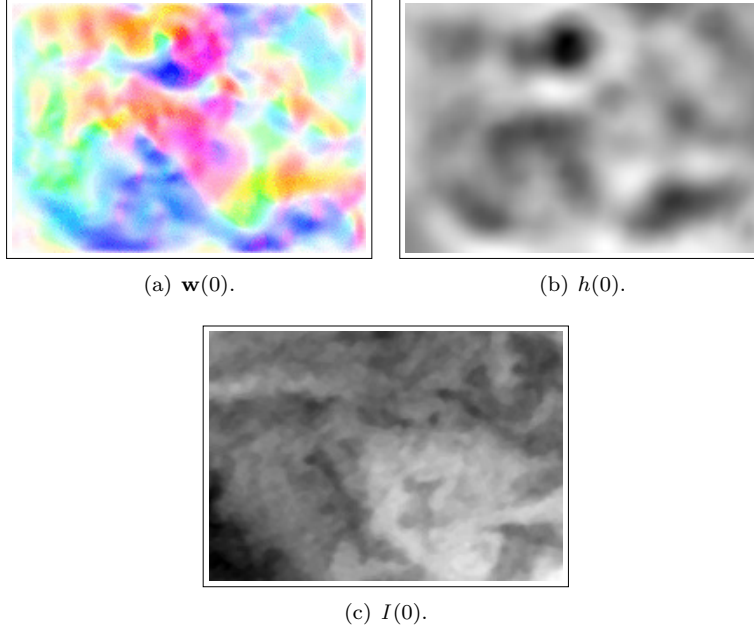
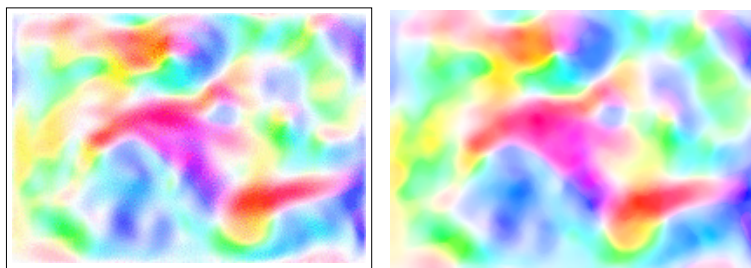


Figure 3: Initial conditions of the simulation run.

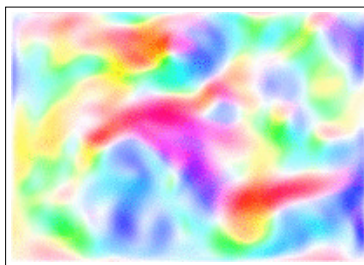
of Eq. (17). The Coriolis force is estimated as  $C = f\mathbf{w}^\perp$ , with  $\mathbf{w}^\perp = \begin{pmatrix} v \\ -u \end{pmatrix}$ , and the gravity force  $G$  is obtained as  $G = \mathbf{a} - C$ . For motion estimated with Sun *et al.* [28], the acceleration is computed as a first-order finite difference and the Coriolis and gravity forces are retrieved with the same equations. Both estimations are compared with the ground truth values  $C_{GT}$  and  $G_{GT}$ , defined by:  $C_{GT} = f\mathbf{w}_{GT}^\perp$  and  $G_{GT} = -g'\nabla h_{GT}$ , where  $\mathbf{w}_{GT}$  and  $h_{GT}$  are the ground truths for motion and mixed layer thickness.

Figures 4 and 5 concern the Coriolis force and show the accuracy of our method: the estimation is close to the ground truth value both in orientation and norm. The displays also provide the results of Sun *et al.* Averaged angular and relative norm errors are respectively of  $6.1^\circ$  and 10% for our method and  $15.9^\circ$  and 39% for Sun *et al.* [28]. This conclusion is not new as this is well known that optical flow methods fail to retrieve an accurate estimation of acceleration and consequently of geophysical forces.

Figures 6 and 7 provide the same displays for the estimation of the gravity force. Statistics on angular error and relative norm errors are  $24.2^\circ$  and 36% for our method against  $56^\circ$  and 38% for Sun *et al.* [28]. Our estimation appears however to be less accurate than the one of the Coriolis force. This comes from the approximation of the shallow water equations and from the removal of friction and diffusion forces in the model.

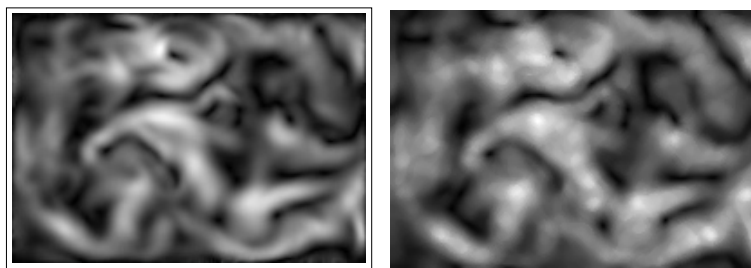


(a) Orientation of  $C$  computed by our method. (b) Orientation of  $C$  computed by Sun *et al.*

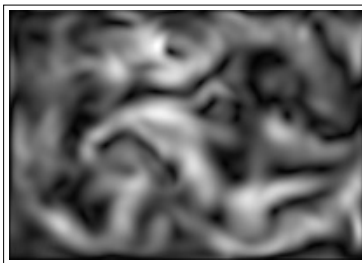


(c) Orientation of  $C_{GT}$ .

Figure 4: Orientation of the Coriolis force.

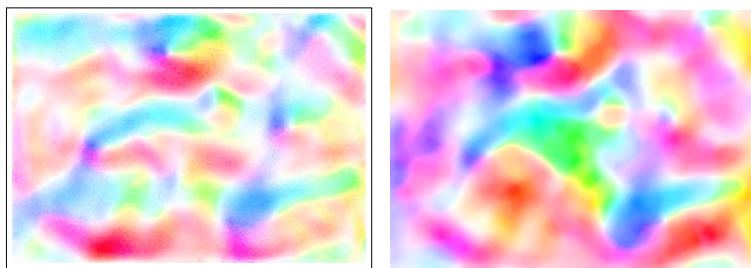


(a) Norm of  $C$  computed by our method. (b) Norm of  $C$  computed by Sun *et al.*

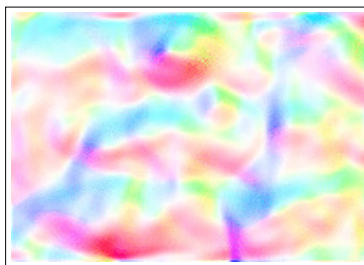


(c) Norm of  $C_{GT}$ .

Figure 5: Norm of the Coriolis force.



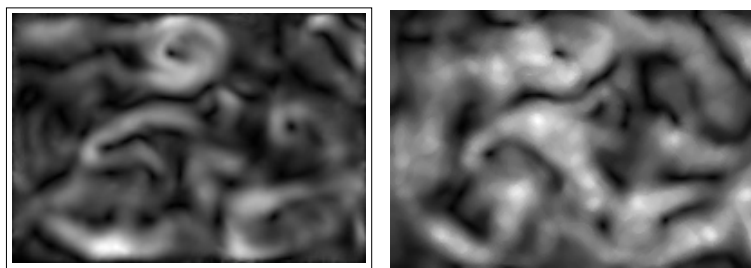
(a) Orientation of  $G$  computed by our method. (b) Orientation of  $G$  computed by Sun *et al.*



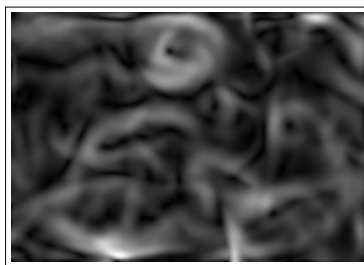
(c) Orientation of  $G_{GT}$ .

Figure 6: Orientation of the gravity force.





(a) Norm of  $G$  computed by our method. (b) Norm of  $G$  computed by Sun *et al.*



(c) Norm of  $G_{GT}$ .

Figure 7: Norm of the gravity force.

## 5.2 Satellite data

After having been quantified in the previous subsection, the method was applied on a database of ocean satellite acquisitions. Two examples are discussed in the following.

A first display, on Figure 8, concerns Sea Surface Temperatures acquired on January 13, 2005 by NOAA-AVHRR. These acquisitions include missing values that are visible as black points or black regions. They mostly originate from sensor failures. They are excluded from the estimation process by giving them a high variance value (an a null impact) in the energy function  $J$  of Eq. (7).

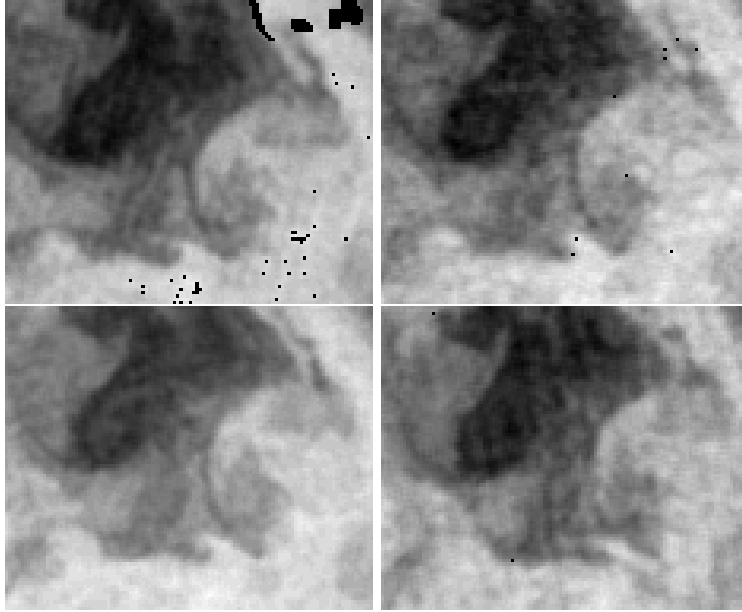


Figure 8: Four SST acquisitions at times  $T_0$ ,  $T_0+3h30$ ,  $T_0+6h30$ ,  $T_0+11h15$ .

The estimation is made with the setting described in Section 4 (see description of  $\mathbf{X}_b$ ,  $B$ ,  $R$ ). The motion results are given at acquisition times on Figure 9, with an arrows display superposed on the image data for a better visualization of the cyclonic motion of the vortex structure.

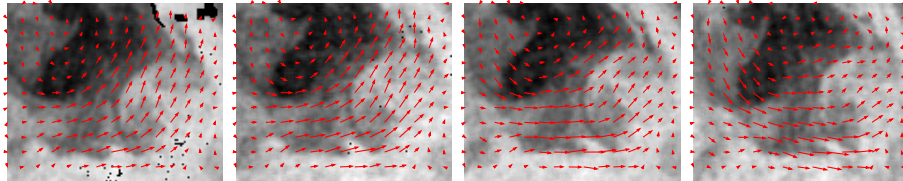


Figure 9: Motion field superposed to the satellite acquisitions.

The estimated acceleration  $\mathbf{a}$  is given on Figure 10, at times corresponding to the acquisition of the second, third and fourth images, with a colors and arrows coding.

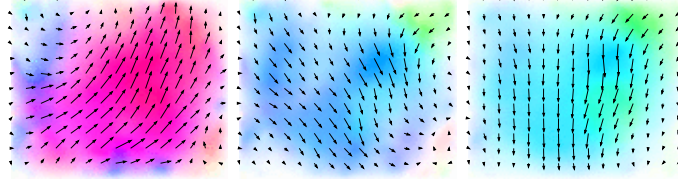


Figure 10: Acceleration at acquisition times.

The Coriolis force  $C$  is then computed, see Figure 11, which is in complete adequacy with the cyclonic mushroom-shaped vortex.

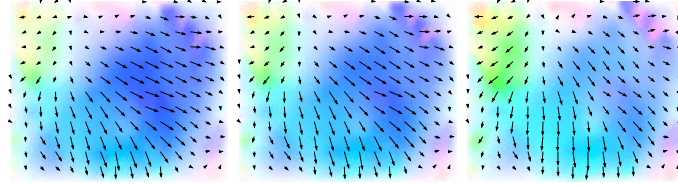


Figure 11: Coriolis force at acquisition times.

Last, Figure 12 displays the estimated gravity force  $G$ .

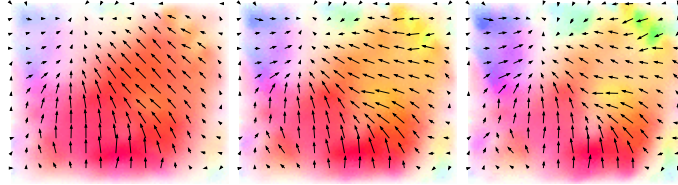


Figure 12: Gravity at acquisition times.

Having knowledge on image and motion fields and on the forces applied to that system allows to use this whole set of information to forecast the future images at short temporal horizon, which is one major application of this work in environmental sciences.

The method is further illustrated with images acquired on April 23, 2007. The satellite acquisitions are displayed on Figure 13. They show an anticyclonic motion field.

The estimated motion fields  $\mathbf{w}(t)$  are visualized on Figure 14, superposed to the acquisitions.

The estimated acceleration  $\mathbf{a}$ , the Coriolis force  $C$  and the gravity  $G$  are respectively given on Figures (15), (16) and (17), with a vectors and colors

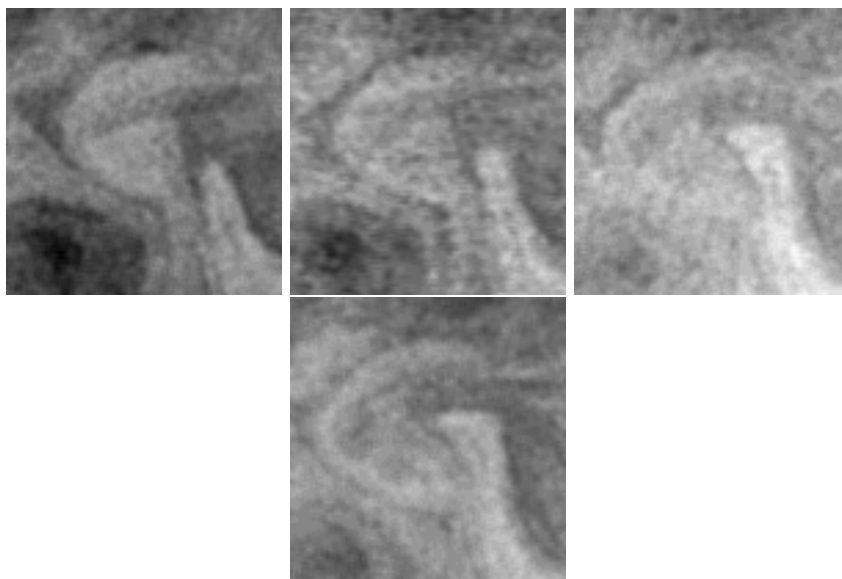


Figure 13: Four Observations (with missing data) at times  $T_0$ ,  $T_0+4h$ ,  $T_0+14h$ ;  $T_0+16h30$ .

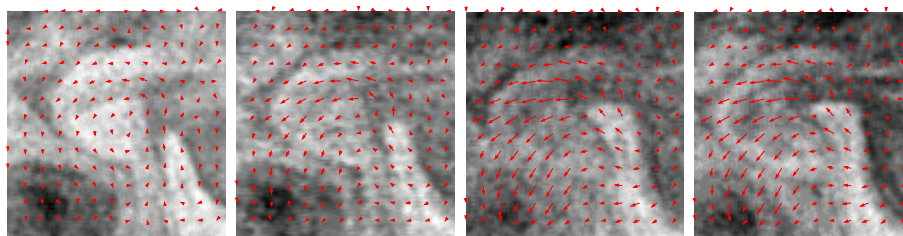


Figure 14: Motion fields superposed to image acquisitions.

coding. We can observe that the map of the Coriolis force is coherent with the anticyclonic vortex, which is evolving on the data.

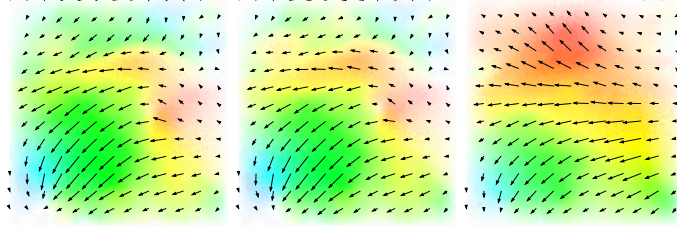


Figure 15: Acceleration  $\mathbf{a}$  at acquisition times.

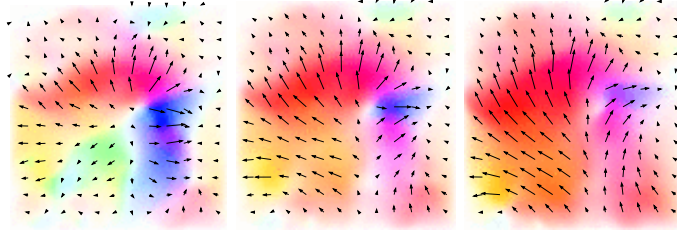


Figure 16: Coriolis force  $C$  at acquisition times.

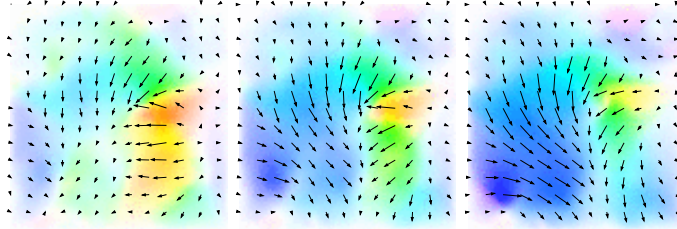


Figure 17: Gravity  $G$  at acquisition times.

### 5.3 Evaluation on short and mid-term forecasting

In this section, we provide experiments for the use of the estimated acceleration for short and mid-term forecasting. This is conducted in an operational context and compared to state-of-the-art results. For providing a quantitative analysis, we make use of synthetic SST images computed by the operational oceanographic software NEMO (Nucleus for European Modeling of the Ocean). NEMO provides an estimation of the ocean state and produces realistic synthetic SST images as reanalysis obtained from data assimilation of satellite acqui-

tions with an ocean physical model. We retrieve<sup>1</sup> a sequence of 456 images (from 2016-01-01 to 2017-04-06) from data acquired over North Atlantic. The space resolution of each pixel is 10 kilometers and the time interval between two images is one day. After computing the reanalyses, we extract a region of interest of size  $481 \times 456$ . Figure 18 displays some of these NEMO SST images.

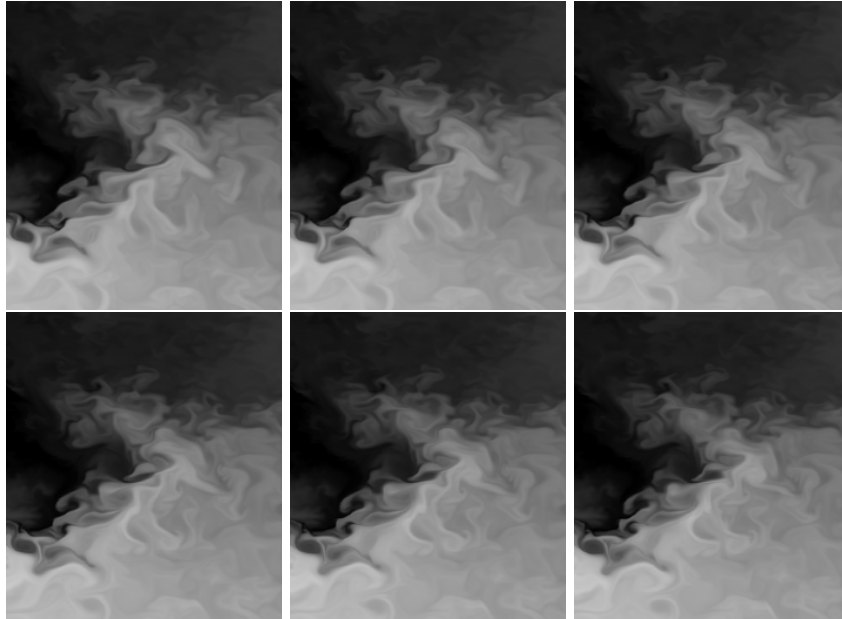


Figure 18: Six pseudo SST observations computed by NEMO.

The sliding window method is designed as follows. We compute a motion field from the first four consecutive synthetic images. Then, from the last image and the corresponding motion field, the state vector is integrated in time using Eqs (2,3) up to the chosen horizon of forecast. The process, estimation and forecast, is iterated on each interval of four consecutive images until having processed the whole sequence.

Our approach is then compared with the competitive Zack *et al.*'s optical flow method [29]. We display on Figure 19 the motion result estimated by our method at observation 4 (see Figure 18) and the result of Zack's method estimated between observations 3 and 4.

Forecasts, at 1-day and 10-day, provided by both motion fields, are compared to the synthetic images produced by NEMO on Figures 20 and 21. We consider the Mean Square Error (MSE) criterion, computed between forecasts and synthetic images, to quantify performances. As it can be seen on Figure 20, short time forecasts obtained by the two motion fields are quite similar. This is

<sup>1</sup><http://marine.copernicus.eu/services-portfolio/access-to-products/>

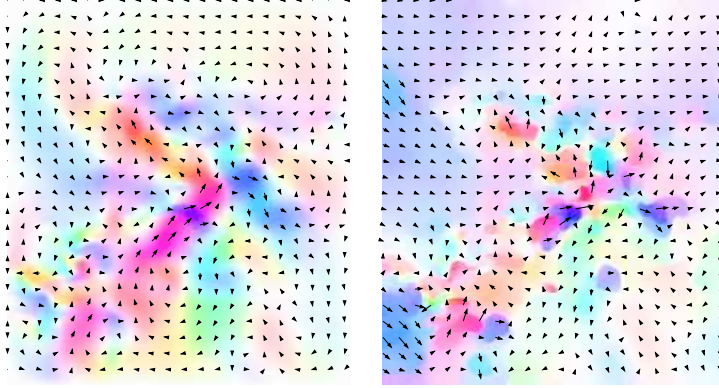


Figure 19: Motion estimation by our method (left) and Zack *et al* (right)

confirmed by the MSE criteria, computed on the whole synthetic SST sequence and displayed on Figure 22. MSE shows a slight advantage for our approach with an average value of 0.093 against 0.102 for Zack *et al*. However, our approach is clearly better than Zack *et al* for larger temporal horizon. This is visible on Figure 23 for 10-days forecasts. The average value of MSE is of 1.88 for our approach against 2.35 for Zack *et al*.

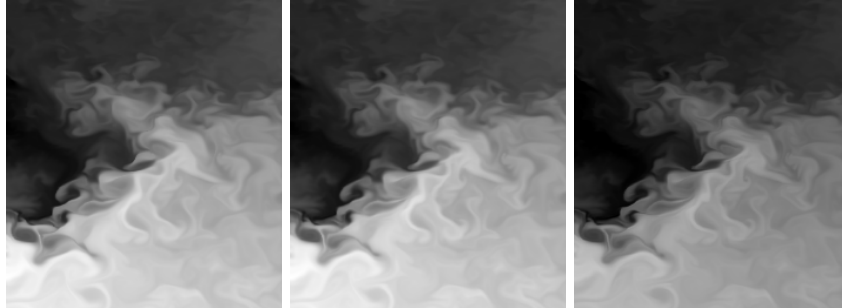


Figure 20: 1-day forecast. Left: our method, middle: Zack *et al*, right: observation.

## 6 Conclusion and Future works

In this paper, a variational approach for assessing the dynamics of a studied system from image data is described. The acquisitions are processed on line with a sliding window approach and the first guess of the acceleration and motion fields on a given window is obtained from the result on the previous one. We design a model of motion and acceleration fields and an energy function, based

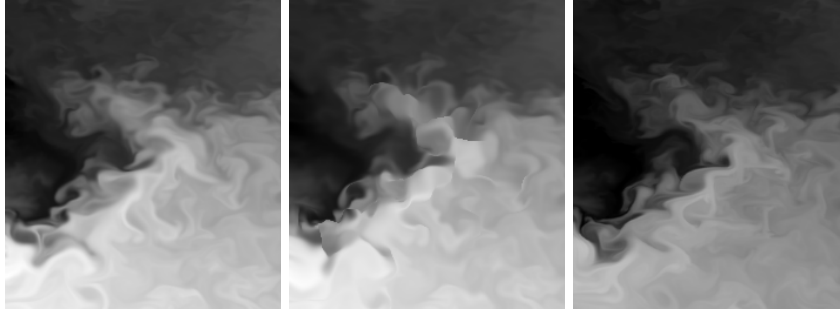


Figure 21: 10-day forecast. Left: our method, middle: Zack *et al*, right: observation.

on space-time constraints, for their estimation. The optimization minimizes the discrepancy between the given images and the ones synthesized by our model. It outputs both the motion field at the beginning of the studied temporal interval and the acceleration field at each time step. The knowledge of this acceleration, as a space-time function, has a strong additional value, compared to its computation by finite difference from the motion fields, as it gives the full knowledge of the physical forces applied on the system.

The experiments conducted on synthetic data quantify the approach and illustrate its potential on fluid-like flows images, such as satellite acquisitions of the atmosphere and ocean.

Future research is two-fold. First, the design of the regularization terms included in the energy function will be reinvestigated, for improving the quantitative accuracy of the estimation, without increasing computational requirements. Second, the whole software will be parallelized with MPI/OpenMP techniques for processing larger size image data and longer duration observation. The constraint of size and duration is the major limitation for an operational use of the method for monitoring local sites. The first foreseen operational application is the forecast of the photovoltaic production at short temporal horizon, through the solar irradiation estimation from image data.

## References

- [1] D. Heitz, E. Mémin, C. Schnörr, Variational fluid flow measurements from image sequences: synopsis and perspectives., *Experiments in fluids* 48 (2010) 369–393. [2](#)
- [2] J. Arnspang, Optic acceleration, in: *International Conference on Computer Vision*, 1988, pp. 364–373. [2](#)
- [3] X. Hu, N. Ahuja, Estimating motion of constant acceleration from image sequences, in: *International Conference on Image Processing*, 1993, pp. 655–659. [2](#)



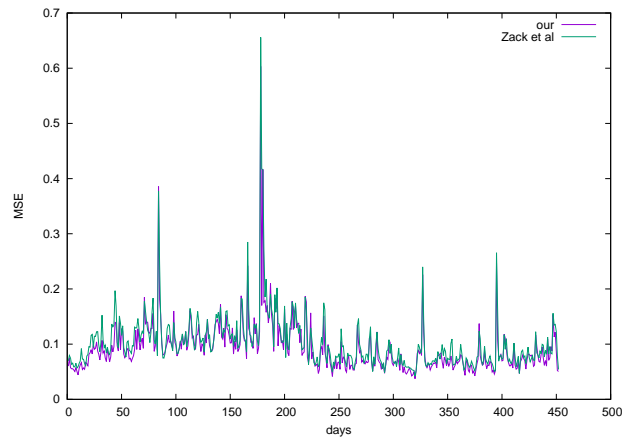


Figure 22: MSE computed on 1-day forecast

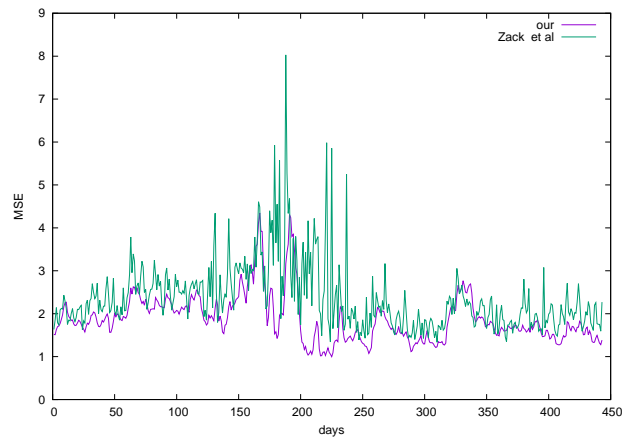


Figure 23: MSE computed on 10-day forecast

- [4] S. Stanković, I. Djurović, R. Herpers, Velocity and acceleration estimation in video sequences by the local polynomial periodogram, in: International Symposium on Signal Processing and Its Applications, 2003. 2
- [5] C. Tomasi, T. Kanade, Shape and motion from image streams under orthography: a factorization method, International Journal on Computer Vision 9 (2) (1992) 137–154. 2
- [6] M. Irani, Multi-frame correspondance estimation using subspace constraints, International Journal on Computer Vision 48 (3) (2002) 173–194. 2
- [7] B. Lucas, T. Kanade, A iterative image registration technique with an application to stereo vision, in: Proc. of 7th International Joint Conference on Artificial Intelligence, Vancouver, Canada, 1981, pp. 674–679. 2
- [8] R. Garg, L. Pizarro, D. Rueckert, L. Agapito, Dense multi-frame optic flow for non-rigid objects using subspace constraints, in: Asian Conference on Computer Vision, Vol. 6495, 2010, pp. 460–473. 2
- [9] S. Ricco, C. Tomasi, Dense Lagrangian motion estimation with occlusions, in: Conference on Computer Vision and Pattern Recognition, 2012, pp. 1800–1807. 2
- [10] R. Garg, A. Roussis, L. Agapito, A variational approach to video registration with subspace constraints, International Journal on Computer Vision 104 (3) (2013) 286–314. 2
- [11] N. Papadakis, E. Mémin, Variational assimilation of fluid motion from image sequence, SIAM Journal on Numerical Analysis 1 (4) (2008) 343–363. 3
- [12] T. Corpetti, P. Héas, E. Mémin, N. Papadakis, Pressure image assimilation for atmospheric motion estimation, Tellus A 61 (1) (2009) 160–178. doi:10.1111/j.1600-0870.2008.00370.x. URL <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1600-0870.2008.00370.x> 3
- [13] O. Titaud, A. Vidard, I. Souopgui, F.-X. Le Dimet, Assimilation of image sequences in numerical models, Tellus Series A : Dynamic meteorology and oceanography 62 (2010) 30–47. doi:10.1111/j.1600-0870.2009.00416.x. 3
- [14] D. Béréziat, I. Herlin, Solving ill-posed image processing problems using data assimilation, Numerical Algorithms 56 (2) (2011) 219–252. 3
- [15] P. Héas, A. Dréneau, C. Herzet, An efficient algorithm for video super-resolution based on a sequential model, Journal on Imaging Sciences 9 (2). 3

- [16] G. Evensen, The ensemble Kalman filter: Theoretical formulation and practical implementation, *Ocean Dynamics* 53 (2003) 343–367. [3](#)
- [17] D. Dee, Bias and data assimilation, *Quarterly Journal of the Royal Meteorological Society* 131 (2005) 3323–3343. [3](#)
- [18] E. Valur Hólm, Lectures notes on assimilation algorithms, Tech. rep., European Centre for Medium-Range Weather Forecasts Reading, U.K (April 2008). [3](#)
- [19] R. H. Byrd, P. Lu, J. Nocedal, A limited memory algorithm for bound constrained optimization, *Journal on Scientific and Statistical Computing* 16 (5) (1995) 1190–1208. [4](#), [6](#)
- [20] A. Apte, C. Jones, A. Stuart, J. Voss, Data assimilation: Mathematical and statistical perspectives, *Int. J. Numer. Meth. Fluids* 56 (2008) 1033–1046. [5](#)
- [21] J.-L. Lions, *Optimal Control of Systems Governed by Partial Differential Equations*, Springer-Verlag, 1971. [6](#)
- [22] I. Herlin, D. Béréziat, N. Mercier, S. Zhuk, Divergence-free motion estimation, in: *European Conference on Computer Vision*, Vol. 7575 of *Lecture Notes on Computer Sciences*, Firenze, Italy, 2012, pp. 15–27. [6](#)
- [23] W. Hundsdorfer, E. Spee, An efficient horizontal advection scheme for the modeling of global transport of constituents, *Monthly Weather Review* 123 (12) (1995) 3,554–3,564. [7](#)
- [24] R. LeVeque, *Numerical Methods for Conservative Laws*, 2nd Edition, *Lectures in Mathematics*, ETH Zürich, Birkhäuser Verlag, 1992. [7](#)
- [25] L. Hascoët, V. Pascual, [The Tapenade Automatic Differentiation tool: Principles, Model, and Specification](#), *ACM Transactions On Mathematical Software* 39 (3).  
URL <http://dx.doi.org/10.1145/2450153.2450158> [7](#)
- [26] I. Herlin, D. Béréziat, E. Huot, Image assimilation and motion estimation of geophysical fluids, Leipzig, Germany, 2012, pp. 2325–2332. [9](#)
- [27] R. Balgovind, A. Dalcher, M. Ghil, E. Kalnay, A stochastic dynamic model for the spatial structure of forecast error statistics, *Monthly Weather Review* 111 (1983) 701–722. [10](#)
- [28] D. Sun, S. Roth, M. Black, Secrets of optical flow estimation and their principles, in: *European Conference on Computer Vision*, 2010, pp. 2432–2439. [11](#), [12](#)
- [29] C. Zack, T. Pock, H. Bischof, A duality based approach for realtime TV-L<sup>1</sup> optical flow, *Lecture Notes in Computer Science* 4713 (2007) 214–223. [21](#)